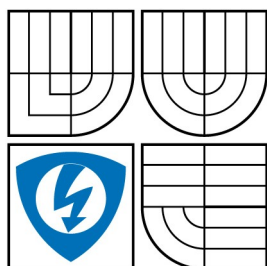


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ  
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH  
TECHNOLOGIÍ  
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
DEPARTMENT OF TELECOMMUNICATIONS

# SKLÁDÁNÍ SNÍMKŮ PANORAMATICKÉHO POHLEDU

PANORAMATIC VIEW RECONSTRUCTION

DIPLOMOVÁ PRÁCE  
MASTER'S THESIS

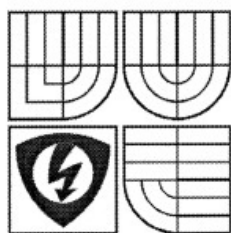
AUTOR PRÁCE  
AUTHOR

Bc. OLDŘICH KUZDAS

VEDOUCÍ PRÁCE  
SUPERVISOR

Ing. KAMIL ŘÍHA, Ph. D.

BRNO 2008



VYSOKÉ UČENÍ  
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

Ústav telekomunikací

# Diplomová práce

magisterský navazující studijní obor  
Telekomunikační a informační technika

**Student:** Kuzdas Oldřich, Bc.

**Ročník:** 2

**ID:** 88579

**Akademický rok:** 2007/08

**NÁZEV TÉMATU:**

## Skládání snímků panoramatického pohledu

### POKYNY PRO VYPRACOVÁNÍ:

Nastudujte teoreticky problematiku sestavení panoramatického obrazu z několika dílčích snímků scény pořízených perspektivní kamerou otáčenou kolem jejího optického středu. Dále nastudujte techniky odstranění nežádoucích ostrých přechodů mezi zdrojovými snímky ve výsledném panoramatickém obraze.

Získané teoretické poznatky využijte při implementaci nastudovaných metod formou samostatné aplikace s možností volby různých testovacích obrazů.

Doporučené softwarové nástroje jsou Visual C++ 2005 Express Edition a knihovny OpenCV pro zpracování obrazu.

### DOPORUČENÁ LITERATURA:

- [1] GONZALEZ R. C., WOODS R. E.: Digital Image Processing, Prentice Hall, New Jersey, 2002
- [2] FISHER R. B.: CVonline: The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision, <http://homepages.inf.ed.ac.uk/rbf/CVonline/>
- [3] PRATA, S.: Mistrovství v C++, Computer Press, Brno 2004

**Termín zadání:** 11.2.2008

**Termín odevzdání:** 28.5.2008

**Vedoucí projektu:** Ing. Kamil Říha, Ph.D.

prof. Ing. Kamil Vrba, CSc.  
předseda oborové rady



### UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

**LICENČNÍ SMLOUVA**  
**POSKYTOVANÁ K VÝKONU PRÁVA UŽÍT ŠKOLNÍ DÍLO**

uzavřená mezi smluvními stranami:

**1. Pan/paní**

Jméno a příjmení: Bc. Oldřich Kuzdas  
Bytem: Halasova 9, 58601, Jihlava  
Narozen/a (datum a místo): 22.1.1984, Jihlava

(dále jen „autor“)

a

**2. Vysoké učení technické v Brně**

Fakulta elektrotechniky a komunikačních technologií  
se sídlem Údolní 244/53, 602 00, Brno  
jejímž jménem jedná na základě písemného pověření děkanem fakulty:  
prof. Ing. Kamil Vrba, CSc.

(dále jen „nabyvatel“)

**Článek 1**  
**Specifikace školního díla**

1. Předmětem této smlouvy je vysokoškolská kvalifikační práce (VŠKP):

- ☐ disertační práce
- ☐ diplomová práce
- ☐ bakalářská práce
- ☐ jiná práce, jejíž druh je specifikován jako .....

(dále jen VŠKP nebo dílo)

Název VŠKP: Skládání snímků panoramatického pohledu

Vedoucí/ školitel VŠKP: Ing. Kamil Říha, Ph.D.

Ústav: Ústav telekomunikací

Datum obhajoby VŠKP: .....

VŠKP odevzdal autor nabyvateli v:

tištěné formě – počet exemplářů .....

elektronické formě – počet exemplářů .....

2. Autor prohlašuje, že vytvořil samostatnou vlastní tvůrčí činností dílo shora popsané a specifikované. Autor dále prohlašuje, že při zpracovávání díla se sám nedostal do rozporu s autorským zákonem a předpisy souvisejícími a že je dílo dílem původním.
3. Dílo je chráněno jako dílo dle autorského zákona v platném znění.
4. Autor potvrzuje, že listinná a elektronická verze díla je identická.

## **Článek 2**

### **Udělení licenčního oprávnění**

1. Autor touto smlouvou poskytuje nabyvateli oprávnění (licenci) k výkonu práva uvedené dílo nevýdělečně užít, archivovat a zpřístupnit ke studijním, výukovým a výzkumným účelům včetně pořizování výpisů, opisů a rozmnoženin.
2. Licence je poskytována celosvětově, pro celou dobu trvání autorských a majetkových práv k dílu.
3. Autor souhlasí se zveřejněním díla v databázi přístupné v mezinárodní síti
  - ☐ ihned po uzavření této smlouvy
  - ☐ 1 rok po uzavření této smlouvy
  - ☐ 3 roky po uzavření této smlouvy
  - ☐ 5 let po uzavření této smlouvy
  - ☐ 10 let po uzavření této smlouvy(z důvodu utajení v něm obsažených informací)
4. Nevýdělečné zveřejňování díla nabyvatelem v souladu s ustanovením § 47b zákona č. 111/ 1998 Sb., v platném znění, nevyžaduje licenci a nabyvatel je k němu povinen a oprávněn ze zákona.

## **Článek 3**

### **Závěrečná ustanovení**

1. Smlouva je sepsána ve třech vyhotoveních s platností originálu, přičemž po jednom vyhotovení obdrží autor a nabyvatel, další vyhotovení je vloženo do VŠKP.
2. Vztahy mezi smluvními stranami vzniklé a neupravené touto smlouvou se řídí autorským zákonem, občanským zákoníkem, vysokoškolským zákonem, zákonem o archivnictví, v platném znění a popř. dalšími právními předpisy.
3. Licenční smlouva byla uzavřena na základě svobodné a pravé vůle smluvních stran, s plným porozuměním jejímu textu i důsledkům, nikoliv v tísní a za nápadně nevýhodných podmínek.
4. Licenční smlouva nabývá platnosti a účinnosti dnem jejího podpisu oběma smluvními stranami.

V Brně dne: .....

.....

Nabyvatel

.....

Autor

## Abstrakt

Tato práce se zabývá krok za krokem procesem sestavování panoramatického obrazu z několika dílčích snímků scény pořízených perspektivní kamerou otáčenou kolem jejího optického středu. Jsou zde popsány algoritmy detekce významných bodů v obraze, možnosti výpočtu matice homografie a metody odstranění nežádoucích ostrých přechodů mezi zdrojovými snímky ve výsledném panoramatickém obraze. Součástí práce je i samostatná aplikace, v níž jsou implementovány některé algoritmy popsané v této práci.

## Klíčová slova

Skládání panoramatického snímků, homografie, perspektivní transformace, SIFT, klíčový bod, RANSAC, lineární prolínání, multipásmové prolínání

## Abstract

This paper deals step by step with process of stitching images taken by perspective camera rotated by its optical center into the panoramic image. There are described keypoint searching algorithms, possibilities of calculating homography matrix and methods of eliminating unwanted seams between source images in final panoramic image. A part of this paper is also standalone application in which are implemented some algorithms described in the work.

## Keywords

Creating panoramic image, homography, perspective transformation, SIFT, keypoint, RANSAC, linear blending, multi-band blending

## Citace

KUZDAS, O. *Skládání snímků panoramatického pohledu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2008. 46 s. Vedoucí diplomové práce Ing. Kamil Říha, Ph.D.

## Prohlášení

Prohlašuji, že svou diplomovou práci na téma „Skládání snímků panoramatického pohledu“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne .....

## Poděkování

Děkuji svému vedoucímu diplomové práce Ing. Kamilu Říhovi, Ph. D. za cenné rady, připomínky a odbornou pomoc, kterou mi poskytl v průběhu řešení.

V Brně dne .....

# Obsah

1. Úvod.....	3
2. Mezisnímkové korespondence.....	4
2.1. Moravcův detektor.....	5
2.2. Harrisův detektor.....	6
2.3. Scale-space.....	8
2.4. SIFT.....	9
2.5. RANSAC.....	12
3. Model perspektivní kamery.....	15
3.1. Perspektivní promítání.....	15
3.2. Matematický popis perspektivní kamery.....	16
3.3. Scéna zachycená dvěma kamerami.....	17
4. Homografie.....	19
4.1. Definice.....	19
4.2. Výpočet homografie.....	21
4.2.1. Bodová korespondence.....	21
4.2.2. Přímková korespondence.....	22
5. Skládání zdrojových snímků.....	25
5.1. Lineární prolínání.....	29
5.2. Multipásmové prolínání.....	29
5.3. Porovnání metod.....	32
5.4. Chyba paralaxe.....	35
6. Vytvořená aplikace.....	36
6.1. Načítání snímků.....	36
6.2. Mezisnímkové korespondence.....	37
6.3. Tvorba panoramatického snímku.....	38
6.4. Zdrojové kódy.....	40
7. Závěr.....	42
8. Použitá literatura.....	44
A. Obsah příloženého CD.....	46

## Seznam obrázků

Obr. 2.1: Model Scale-space.....	9
Obr. 2.2: Sestrojení Scale-space za pomoci rozdílů Gaussových funkcí.....	10
Obr. 2.3: Definice sousedních bodů pro hledání extrémů.....	11
Obr. 2.4: Deskriptor významného bodu dle metody SIFT.....	12
Obr. 2.5: Zjištění mezisnímkové korespondence bez použití RANSAC.....	13
Obr. 2.6: Zjištění mezisnímkové korespondence s použitím RANSAC.....	14
Obr. 3.1: Perspektivní projekce.....	15
Obr. 3.2: Scéna zachycená dvěma kamerami.....	18
Obr. 4.1: Perspektivní transformace bodu.....	19
Obr. 5.1: Zdrojové snímky.....	25
Obr. 5.2: Transformované snímky před spojením do panoramatického obrazu.....	25
Obr. 5.3: Výsledné panorama sestavené přeložením jednoho snímku přes druhý.....	26
Obr. 5.4: Váhová matice $\mathbf{W}$ .....	27
Obr. 5.5: Váhová matice $\mathbf{W}_{\text{MAX}}$ .....	27
Obr. 5.6: Výsledný panoramatický snímek dle obr. 5.5.....	28
Obr. 5.7: Lineární prolínání transformovaných zdrojových snímků.....	29
Obr. 5.8: Výsledné panorama s použitím lineárního prolnutí.....	29
Obr. 5.9: Multipásmové prolínání, výsledný panoramatický snímek.....	31
Obr. 5.10: Multipásmové prolínání, $k = 1$ .....	32
Obr. 5.11: Multipásmové prolínání, $k = 2$ .....	32
Obr. 5.12: Multipásmové prolínání, $k = 3$ .....	32
Obr. 5.13: Panoramatický snímek 1.....	33
Obr. 5.14: Panoramatický snímek 2.....	33
Obr. 5.15: Panoramatický snímek 3.....	34
Obr. 5.16: Panoramatický snímek 4.....	34
Obr. 5.17: Chyba paralaxe.....	35
Obr. 6.1: Načítání zdrojových snímků.....	36
Obr. 6.2: Mezisnímkové korespondence.....	37
Obr. 6.3: Tvorba panoramatického snímku.....	40



# 1. Úvod

I když jsou v dnešní době k dispozici na trhu digitální fotoaparáty s rozlišením přes 10 megapixelů, stále to pro určité účely není dost. Jako příklad lze zmínit velkoformátový tisk nebo interaktivní aplikace, kdy si lze na snímku přiblížit i vzdálené objekty scény ve vysoké kvalitě. A právě zde přichází na řadu skládání snímků panoramatického pohledu.

Proces skládání snímků panoramatického pohledu sestává z několika základních kroků, které na sebe ovšem pevně navazují. Těmito kroky jsou zjištění korespondencí mezi zdrojovými snímky, výpočet matice perspektivní transformace, samotná perspektivní transformace jednotlivých zdrojových snímků a posledním krokem je spojení transformovaných zdrojových snímků v jeden snímek panoramatický. Vždy se volí jeden zdrojový snímek jako hlavní, což znamená, že zobrazovací rovina tohoto snímku bude zároveň zobrazovací rovinou výsledného panoramatického snímku. Pro získání kvalitní panoramatické fotografie je zapotřebí, aby se scény zachycené na zdrojových snímcích částečně překrývaly. Minimální míra překrytí by se měla být 20%.

Stejně jako samotný proces skládání snímků panoramatického pohledu lze i tuto práci rozdělit na čtyři hlavní části. První část se věnuje možnostem získávání korespondencí mezi zdrojovými snímky, následují možnosti výpočtu matice perspektivní transformace a popis metod pro odstranění nežádoucích ostrých přechodů mezi zdrojovými snímky ve výsledném panoramatickém obraze. Závěr práce obsahuje popis vytvořené aplikace, ve které jsou implementovány některé níže popsané algoritmy.

## 2. Mezisnímkové korespondence

Korespondující body dvou dílčích zdrojových snímků je možné hledat manuálně nebo za pomoci speciálních algoritmů.

Možnost manuálního zadávání mezisnímkových korespondencí sice nabízí plnou kontrolu nad jejich vytvářením, ale je nutné zvážit komplikace, které s sebou tato možnost přináší. Prvním potenciálním problémem je přesnost s jakou musí být voleny korespondující body mezi zdrojovými snímky. I s drobnými odchylkami mezi zvolenými a skutečnými mezisnímkovými korespondencemi lze sice úspěšně složit výsledný panoramatický snímek, ovšem snímek bude vlivem odchylek deformovaný. Dále musíme uvažovat časovou náročnost, která je přímo úměrná počtu zdrojových snímků, ze kterých se výsledný panoramatický snímek bude skládat. Sestává-li budoucí panoramatický snímek z několika desítek zdrojových snímků, může manuální získávání mezisnímkových korespondencí zabrat i několik hodin

Vhodnou alternativou k manuálnímu získávání mezisnímkových korespondencí je využití některého z algoritmů, které tyto korespondence dokáží najít automaticky, relativně přesně a za zlomek času. Schválně zdůrazňuji relativně, protože jejich přesnost je ovlivněna správnou volbou klíčových parametrů vyhledávacích algoritmů. Proto je vhodné vygenerované body ještě překontrolovat a popřípadě upravit jejich polohu nebo upravit parametry algoritmu pro vyhledávání mezisnímkových korespondencí a použít jej znovu..

Metoda vyhledávání korespondencí mezi zdrojovými snímky, kdy by se pro každý obrazový bod jednoho zdrojového snímku hledal odpovídající bod ve druhém zdrojovém snímku, je výpočetně velmi náročná. I když se tomuto postupu nelze vyhnout, lze jeho výpočetní náročnost významně snížit, a to snížením počtu obrazovým bodů ve zdrojových snímcích, jenž se budou mezi sebou porovnávat. Optimalizace spočívá v nalezení tzv. významných (klíčových) bodů ve všech zdrojových snímcích výsledného panoramatického snímku. Ke každému významnému bodu je také připojen deskriptor, což je definovaný popis jeho okolí. Deskriptor může být například vektor jasů okolních obrazových bodů nebo, v případě metody SIFT, histogram gradientů okolí významného bodu. Jsou-li ve všech

zdrojových snímcích definovány významné body spolu s jejich deskriptory, lze porovnáváním deskriptorů významných bodů dvou různých zdrojových snímků získat mezisnímkové korespondence. Významné body jsou označeny jako korespondující, pokud se jejich deskriptory liší maximálně o přesně definovanou odchylku. Je-li hodnota této odchylky zvolena příliš nízká, může být počet mezisnímkových korespondencí nižší než je přípustná mez pro výpočet matice perspektivní transformace. Naopak, je-li hodnota této odchylky zvolena zase příliš vysoká, mohou být některé mezisnímkové korespondence chybně označeny, což znamená, že významné body označené jako korespondující si, se nacházejí v neodpovídajících si oblastech scény. K eliminaci chybně označených mezisnímkových korespondencí se používá algoritmus RANSAC.

## 2.1. Moravcův detektor

Jeden z prvních algoritmů zabývajících se problematikou hledání významných bodů v obraze je Moravcův algoritmus vyvinutý Hans P. Moravcem roku 1977.

V Moravcově algoritmu se pro každý obrazový bod ve zdrojovém snímku vypočítává jasová odlišnost okna vytvořeného okolo zkoumaného bodu od okna posunutého o  $(u, v)$ , kde

$$(u, v) \in \{(-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 0), (0, 1), (1, -1), (1, 0), (1, 1)\} \quad , \quad (2.1)$$

podle

$$E(u, v) = \sum_{\forall x, y} \{ [I(x+u, y+v) - I(x, y)]^2 \} \quad , \quad (2.2)$$

kde  $E$  je matice rozdílů jasů okna posunutého o  $(u, v)$  od okna vytvořeného okolo zkoumaného bodu o souřadnicích  $(x, y)$  a  $I$  je jasová matice zdrojového snímku.

Následně je ze všech  $E(u, v)$  pro daný bod vybrána ta s nejnižší hodnotou  $E_{\min}$ , která je porovnávána s prahem  $c$ . Pokud platí vztah (2.3), pak je zkoumaný bod označen jako významný.

$$E_{\min} > c \quad . \quad (2.3)$$

Největší výhodou tohoto algoritmu je jeho jednoduchost, s čímž souvisí výpočetní nenáročnost. Na druhou stranu má Moravcův algoritmus několik základních nedostatků, mezi které patří náchylnost na šum, nízká odolnost vůči rotaci nebo změně měřítka či jasů zdrojových snímků a také nízká stabilita vygenerovaný významných bodů.

Na základě principu Moravcova algoritmu vznikl algoritmus Harris Stephnes, který odstraňuje některé nedostatky svého předchůdce, jako je například malá odolnost vůči šumu. Zároveň se vyznačuje výrazně vyšší stabilitou nalezených významných bodů.

## 2.2. Harrisův detektor

U zrodu tohoto detektoru stáli pánové Chris Harris a Mike Stephens. Harrisův detektor využívá na rozdíl od Moravcova detektoru kruhové okénko s Gaussovým rozložením vnitřních hodnot, které způsobí, že s rostoucí vzdáleností bodu od zkoumaného bodu, klesá význam jeho hodnoty.

Základem Harrisova detektoru je výpočet autokorelační matice  $\mathbf{A}$  pro každý obrazový bod snímku dle

$$\mathbf{A}(x, y) = \sum_{(u, v) \in \mathbf{W}(x, y)} \mathbf{W}(u, v) [\mathbf{I}(u, v) - \mathbf{I}(u + \Delta x, v + \Delta y)]^2, \quad (2.4)$$

kde  $\mathbf{W}$  představuje okno vytvořeného okolo obrazového bodu zdrojového snímku na souřadnicích  $(x, y)$  s Gaussovým rozložením vnitřních hodnot,  $\mathbf{I}$  je matice jasových hodnot zdrojového snímku.

Pomocí Taylorova (viz [13]) rozvoje můžeme odčítanec z rozdílu jasových hodnot ze vztahu (2.4) upravit jako

$$\mathbf{I}(u + \Delta x, v + \Delta y) \approx \mathbf{I}(u, v) + \mathbf{I}_x(u, v) \Delta x + \mathbf{I}_y(u, v) \Delta y, \quad (2.5)$$

$$\mathbf{I}(u, v) + \mathbf{I}_x(u, v) \Delta x + \mathbf{I}_y(u, v) \Delta y = \mathbf{I}(u, v) + [\mathbf{I}_x(u, v), \mathbf{I}_y(u, v)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \quad (2.6)$$

kde  $\mathbf{I}_x$  a  $\mathbf{I}_y$  jsou parciální derivace jasové matice  $\mathbf{I}$ . Dosazením upraveného vztahu z (2.6) do (2.4) získáme

$$\mathbf{A}(x, y) \approx \sum_{(u, v) \in \mathbf{W}(x, y)} \left( [\mathbf{I}_x(u, v), \mathbf{I}_y(u, v)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 = [\Delta x, \Delta y] \mathbf{M}(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \quad (2.7)$$

kde

$$\mathbf{M}(x, y) = \sum_{(u, v) \in \mathbf{W}(x, y)} \begin{bmatrix} \mathbf{I}_x(u, v)^2 & \mathbf{I}_x(u, v) \mathbf{I}_y(u, v) \\ \mathbf{I}_x(u, v) \mathbf{I}_y(u, v) & \mathbf{I}_y(u, v)^2 \end{bmatrix}, \quad (2.8)$$

$$\mathbf{M}(x, y) = \begin{bmatrix} \sum_{(u, v) \in \mathbf{W}(x, y)} \mathbf{I}_x(u, v)^2 & \sum_{(u, v) \in \mathbf{W}(x, y)} \mathbf{I}_x(u, v) \mathbf{I}_y(u, v) \\ \sum_{(u, v) \in \mathbf{W}(x, y)} \mathbf{I}_x(u, v) \mathbf{I}_y(u, v) & \sum_{(u, v) \in \mathbf{W}(x, y)} \mathbf{I}_y(u, v)^2 \end{bmatrix}. \quad (2.9)$$

Vyjádříme-li si (2.9) symbolicky, pak dostaneme

$$\mathbf{M}(x, y) = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{C} \end{bmatrix}. \quad (2.10)$$

Autoři detektoru definovali pro výpočet mapy rohovosti následující rovnici

$$\mathbf{R}(x, y) = (\mathbf{AC} - \mathbf{B}^2) - k(\mathbf{A} - \mathbf{C})^2, \quad (2.11)$$

kde je  $\mathbf{R}$  je mapa rohovitosti a  $k$  je konstanta zjištěná experimentálně. Nejlepších výsledků metoda dosahuje při  $k \in \langle 0,04; 0,06 \rangle$ . Hodnoty v  $\mathbf{R}$  nižší než definovaná prahová hodnota se nulují. Body s nenulovým parametrem rohovosti jsou označeny jako významné.

Oba výše zmíněné algoritmy lze použít pouze tehdy, pokud mají zdrojové snímky stejné měřítko. Proto byl vyvinut algoritmus SIFT, který převádí informace ze snímků do měřítkově nezávislých interpretací, na základě kterých definuje polohu a deskriptor významných bodů ve zdrojových snímcích. Pro lepší pochopení algoritmu SIFT je nutné si nejdříve popsat algoritmus Scale-space, ze kterého SIFT algoritmus vychází.

## 2.3. Scale-space

Scale-space algoritmus představuje převod obrazových dat z prostorového vyjádření na vyjádření pomocí kmitočtového spektra. Obrazová data jsou reprezentována různými kmitočty. Vyšší kmitočty nesou informace o detailech snímku a naopak nízké kmitočty nesou informace o konturách objektů zachycených na snímku. Volbou vhodné filtrační funkce lze ze zdrojových snímků získat požadovaná data.

Pro převod obrazových dat do Scale-space vyjádření se používá Gaussovska funkce, která se vůči obrazovým datům chová jako dolní propust, a mezní kmitočet je nastavován pomocí parametru  $\sigma$ .

Nejprve si definujeme Gaussovu funkci pro diskretní obraz ve 2D zobrazovací rovině

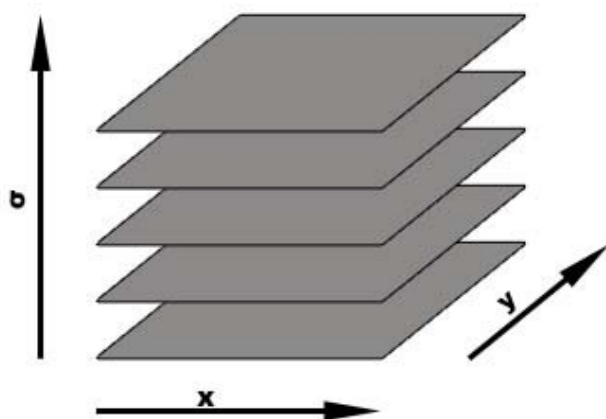
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \quad (2.12)$$

kde  $x$  a  $y$  představují souřadnice bodu v obraze a  $\sigma$  je „měřítkový parametr“.

Nyní si můžeme definovat vrstvu Scale-space

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \quad (2.13)$$

kde  $*$  značí konvoluci,  $L(\sigma)$  je vrstva Scale-space,  $G(\sigma)$  je Gaussova funkce a  $I$  představuje matici jasových hodnot transformovaného obrazu.



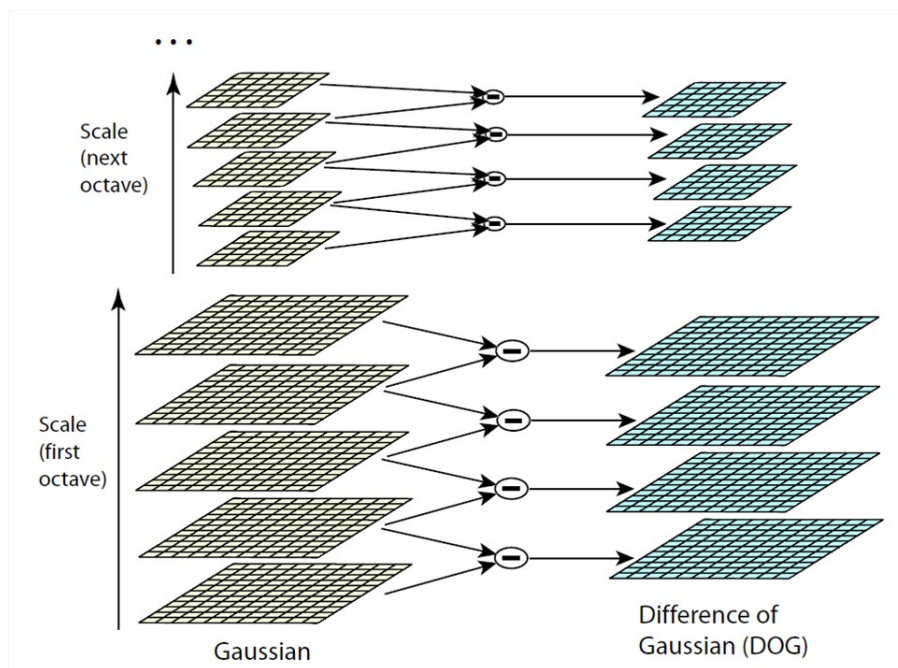
Obr. 2.1: Model Scale-space

## 2.4. SIFT

Scale invariant feature transformation transformuje informace ze snímku do měřítkově nezávislých interpretací. Tato metoda je podrobně popsána ve článku [8], který napsal sám autor metody David Lowe. Získané významné body a jejich deskriptory jsou nezávislé na rotaci a změně měřítka či jasu.

Algoritmus SIFT se skládá z těchto kroků:

1. Sestrojení Space-scale pro všechna měřítka pomocí rozdílů Gaussových funkcí (viz obr. 2.2).
2. Nalezení lokálních extrémů.
3. Přiřazení orientací významným bodům.
4. Výpočet deskriptorů významných bodů.

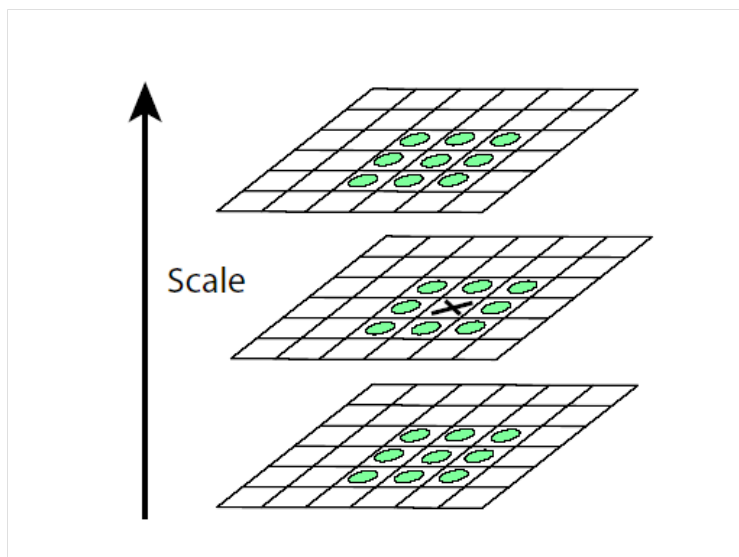


**Obr. 2.2:** Sestrojení Scale-space za pomoci rozdílu Gaussových funkcí (převzato z [8], str. 6)

Počet vrstev DoG má přímý vliv na počet nalezených významných bodů. Použití více vrstev znamená nalezení více významných bodů, ale za cenu vyšších výpočetních nároků a také takto získané významné body mají i nižší odolnost vůči geometrickým a jasovým změnám zdrojového obrazu. Sám autor metody stanovil experimentální formou doporučený počet vrstev DoG na 3. Po vygenerování Scale-space lze přejít k hledání lokálních extrémů.

Hledání lokálních extrémů probíhá ve všech měřících pro všechny body, které mají všech 26 sousedních bodů definovaných dle obr. 2.3. Za významné body jsou považovány ty, jejichž hodnota je menší nebo naopak větší než hodnota všech ostatních sousedních bodů.





**Obr. 2.3:** Definice sousedních bodů pro hledání extrémů (převzato z [8], str. 7)

Ve třetí fázi je významným bodům získaným v předchozím kroku přiřazena jedna nebo více orientací vypočtených z jejich okolí. Kvůli zachování měřítkové nezávislosti se při výpočtech používá plocha dle měřítka daného bodu. Tyto plochy byly již vypočítány během předchozího kroku, tudíž není nutné je počítat znovu. Pro každý bod této plochy se vypočítá velikost gradientu (2.14) a orientace (2.15).

$$\mathbf{m}(x, y) = \sqrt{[\mathbf{I}(x+1, y) - \mathbf{I}(x-1, y)]^2 + [\mathbf{I}(x, y+1) - \mathbf{I}(x, y-1)]^2} \quad , \quad (2.14)$$

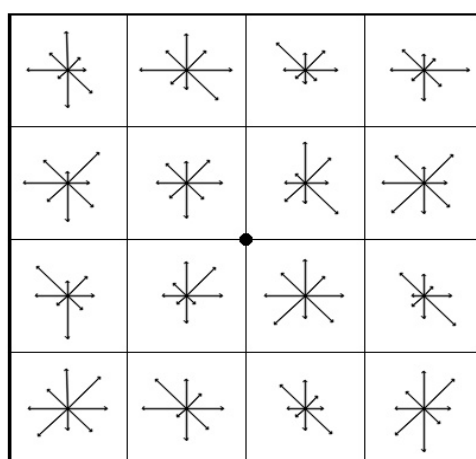
$$\boldsymbol{\theta}(x, y) = \arctan\left(\frac{\mathbf{I}(x, y+1) - \mathbf{I}(x, y-1)}{\mathbf{I}(x+1, y) - \mathbf{I}(x-1, y)}\right) \quad , \quad (2.15)$$

kde  $x$  a  $y$  jsou souřadnice zkoumaného bodu,  $\mathbf{m}$  je matice gradientů,  $\boldsymbol{\theta}$  je matice odpovídajících orientací a  $\mathbf{I}$  představuje jasovou matici měřítka zkoumaného bodu.

Pro každý významný bod je poté sestaven histogram orientací vypočtený z gradientů v jeho okolí. Histogram je popsán 36-ti biny pokrývajícími celých 360°. K histogramu je také připočítán gradient samotného významného bodu, ovšem po přepočítání pomocí okénka s Gaussovým rozložením při  $\sigma = 1,5 m$ , kde  $m$  představuje měřítko významného bodu. Orientace gradientu s nejvyšší hodnotou představuje výslednou orientaci významného bodu. Existují-li v histogramu další gradienty, jejichž hodnoty jsou větší než 80% hodnoty maximálního gradientu, jsou vytvořeny nové významné body se stejnými souřadnicemi jako

původní bod, ovšem s jinými orientacemi. Ikdyž se bodů s více orientacemi vyskytuje velmi málo (asi 15%), představují významné zvýšení stability nalezených významných bodů.

Posledním krokem je tvorba deskriptorů. Deskriptor představuje popis okolí významného bodu, na jehož základě se hledají mezisnímkové korespondence. Stejně jako v předchozím kroku je nutná znalost gradientů v okolí významného bodu, lze tedy využít již vypočtených hodnot, čímž lze ušetřit výpočetní výkon. Metoda SIFT rozděluje okolí významných bodů na množinu  $4 \times 4$  suboblastí a v každé z nich vytváří vlastní histogram sestávající z osmi binů pokrývajících celých  $360^\circ$ . Každý významný bod je tedy popsán deskriptorem, který má  $4 \times 4 \times 8 = 128$  hodnot.



**Obr. 2.4:** Deskriptor významného bodu dle metody SIFT

Máme-li nalezeny významné body pro všechny zdrojové snímky, můžeme přejít k poslednímu kroku, čímž je nalezení párů odpovídajících si významných bodů dvou různých zdrojových snímků. Mezisnímkové korespondence získáme porovnáním deskriptorů všech významných bodů jednoho snímku s deskriptory všech významných bodů jiného zdrojového snímku. Dva významné body různých snímků jsou označeny jako korespondující, pouze pokud se jejich deskriptory liší maximálně o přesně definovanou odchylku.

## 2.5. RANSAC

Mohou nastat dva nežádoucí stavy vlivem nevhodně zvoleného prahu pro určení mezisnímkových korespondencí. První situace nastává, pokud je práh zvolen příliš nízko.

Počet nalezených mezisnímkových korespondencí je nižší než je minimální počet nutný pro možnost výpočtu matice perspektivní transformace. V tomto případě nezbyvá než upravit hodnotu prahu pro určení mezisnímkových korespondencí a hledání opakovat. Na druhou stranu, je-li práh zvolen příliš vysoko, budou jako korespondující si body označeny i významné body nacházející se v neodpovídajících si oblastech scény. Tuto situaci lze řešit bez nutnosti opakovaného hledání mezisnímkových korespondencí. Řešením je aplikování metody RANSAC.

Metoda RANSAC se skládá z těchto kroků:

1. Ze všech zjištěných mezisnímkových korespondencí jsou náhodně vybrány čtyři páry korespondujících bodů, ze kterých je vypočítána matice homografie **H**.
2. Všechny mezisnímkové korespondence jsou přepočítány za pomoci této vypočtené matice homografie **H**, zda si skutečně odpovídají. Kvůli celočíselnému vyjádření souřadnic mezisnímkových korespondencí může dojít vlivem zaokrouhlování ke špatné klasifikaci chybných korespondencí, proto se zde pracuje s určitou předem definovanou tolerancí.
3. Body 1. a 2. se  $n$ -krát opakují, kde  $n$  je definovaný počet iterací. Hledá se taková iterace, která správně transformuje největší počet mezisnímkových korespondencí.



Obr. 2.5: Zjištěné mezisnímkové korespondence bez použití RANSAC

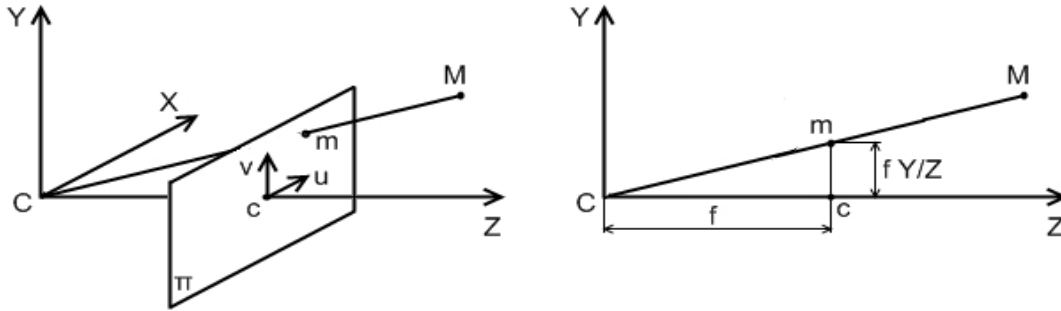


**Obr. 2.6:** Zjištěné mezníkové korespondence s použitím RANSAC

### 3. Model perspektivní kamery

Model perspektivní kamery popisuje způsob promítání bodu ze 3D souřadného systému do 2D projektivní roviny.

#### 3.1. Perspektivní promítání



Obr. 3.1: Perspektivní projekce (vlevo 3D prostor, vpravo 2D řez)

Položíme-li optický střed kamery  $C$  (viz obr. 3.1) do počátku 3D souřadného systému a definujeme-li projektivní rovinu  $\pi$  kolmou na osu  $Z$  3D souřadného systému, která je ve vzdálenosti  $f$  od bodu  $C$ , pak bude bod  $\mathbf{M} = [X, Y, Z]^T$  perspektivně promítán do projektivní roviny  $\pi$  jako bod  $\mathbf{m} = [u, v]^T$ , což lze pomocí homogenních souřadnic obou bodů vyjádřit jako

$$\tilde{\mathbf{m}} \simeq \mathbf{P} \tilde{\mathbf{M}} \quad , \quad (3.1)$$

kde  $\mathbf{P}$  je transformační matice perspektivního promítání. Rozepsáním a úpravou rovnice (3.1) dostaneme

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \simeq \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = Z \begin{bmatrix} f \frac{X}{Z} \\ f \frac{Y}{Z} \\ 1 \end{bmatrix} \quad , \quad (3.2)$$

$$Z \tilde{\mathbf{m}} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{M} \quad , \quad (3.3)$$

kde rovnice (3.3) představuje rovnici perspektivního promítání bodu  $\mathbf{M}$  z 3D souřadného systému do 2D projektivní roviny, přičemž je poloha bodu  $\mathbf{m}$  v 2D projektivní rovině vyjádřena pomocí homogenních souřadnic.

### 3.2. Matematický popis perspektivní kamery

V kapitole 3.1 jsme si popsali situaci, kdy je optický střed kamery  $\mathbf{C}$  totožný se počátkem 3D souřadného systému a projektivní rovina  $\pi$  je kolmá na osu  $Z$  3D souřadného systému. Pro případy, kdy nejsou tyto základní podmínky splněny, je nutné definici transformační matice perspektivního promítání  $\mathbf{P}$  doplnit na

$$\mathbf{P} = \mathbf{A} \mathbf{P}_0 \mathbf{G} , \quad (3.4)$$

$$\mathbf{P}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} , \quad (3.5)$$

kde  $\mathbf{A}$  je matice intrinsických parametrů a  $\mathbf{G}$  představuje extrinsické parametry.

Intrinsické parametry kamery vyjadřují parametry zkreslení obrazu zachycené scény perspektivní kamerou a jsou popsány jako

$$\mathbf{A} = \begin{bmatrix} s_u & 0 & u_0 \\ 0 & s_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} , \quad (3.6)$$

kde  $s_u$  a  $s_v$  určují měřítko obou os projektivní roviny  $\pi$  a parametry  $u_0$  a  $v_0$  udávají polohu počátku souřadného systému projektivní roviny  $\pi$ , tento počátek je na obr. 3.1 zachycen jako  $\mathbf{c}$ .

Mezi extrinsické parametry kamery patří poloha optického středu kamery  $\mathbf{C}$  vůči počátku 3D souřadného systému. Tato poloha je vyjádřena pomocí tříprvkového vektoru  $\mathbf{t}$ . Dále mezi extrinsické parametry kamery patří rotace kamery vůči 3D souřadnému systému, což je popsáno maticí  $\mathbf{R}$  o rozměrech  $3 \times 3$ . Matici  $\mathbf{G}$  lze tedy rozepsat jako

$$\mathbf{G} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_X \\ r_{21} & r_{22} & r_{23} & t_Y \\ r_{31} & r_{32} & r_{33} & t_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix} . \quad (3.7)$$

Dosazením rovnice (3.4) do (3.1) dostaneme

$$\tilde{\mathbf{m}} \simeq \mathbf{A} \mathbf{P}_0 \mathbf{G} \tilde{\mathbf{M}} , \quad (3.8)$$

což lze postupně upravit dle

$$\tilde{\mathbf{m}} \simeq \mathbf{A} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_X \\ r_{21} & r_{22} & r_{23} & t_Y \\ r_{31} & r_{32} & r_{33} & t_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} , \quad (3.9)$$

$$\tilde{\mathbf{m}} \simeq \mathbf{A} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11}X + r_{12}Y + r_{13}Z + t_X \\ r_{21}X + r_{22}Y + r_{23}Z + t_Y \\ r_{31}X + r_{32}Y + r_{33}Z + t_Z \\ 1 \end{bmatrix} , \quad (3.10)$$

$$\tilde{\mathbf{m}} \simeq \mathbf{A} \begin{bmatrix} r_{11}X + r_{12}Y + r_{13}Z + t_X \\ r_{21}X + r_{22}Y + r_{23}Z + t_Y \\ r_{31}X + r_{32}Y + r_{33}Z + t_Z \end{bmatrix} = \mathbf{A} [\mathbf{R} \mathbf{M} + \mathbf{t}] . \quad (3.11)$$

Rovnici (3.11) lze pomocí úplné rovnosti vyjádřit jako

$$\mathbf{Z} \tilde{\mathbf{m}} = \mathbf{A} [\mathbf{R} \mathbf{M} + \mathbf{t}] . \quad (3.12)$$

### 3.3. Scéna zachycená dvěmi kamerami

Umístíme-li optický střed kamery  $\mathbf{C}$  do počátku 3D souřadného systému a bude-li optický střed druhé kamery  $\mathbf{C}'$  vůči němu posunut dle vektoru  $\mathbf{t}$ , pak můžeme vyjádřit body  $\mathbf{m}$  a  $\mathbf{m}'$  jako

$$\mathbf{Z} \tilde{\mathbf{m}} = \mathbf{A} \mathbf{M} , \quad (3.13)$$

$$\mathbf{Z}' \tilde{\mathbf{m}}' = \mathbf{A}' \mathbf{R} \mathbf{M} + \mathbf{A}' \mathbf{t} \quad (3.14)$$

Úpravou rovnice (3.13) dostaneme

$$\mathbf{M} = \mathbf{Z} \mathbf{A}^{-1} \tilde{\mathbf{m}} \quad (3.15)$$

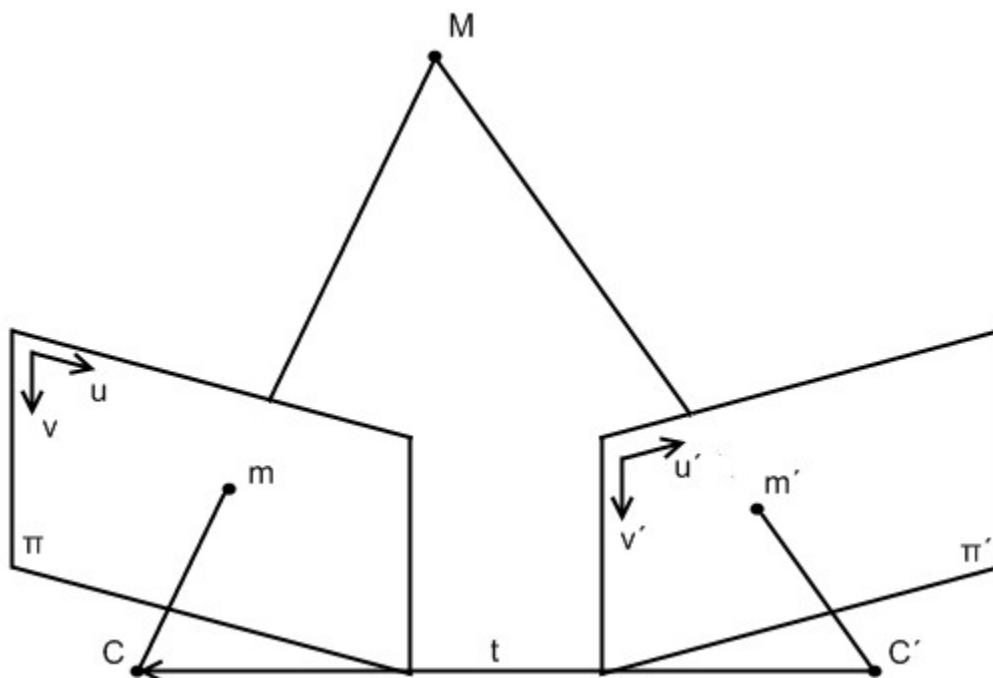
kde  $\mathbf{A}^{-1}$  představuje inverzní matici intrinsických parametrů. Pro inverzní matici platí

$$\mathbf{A}^{-1} \mathbf{A} = \mathbf{A} \mathbf{A}^{-1} = \mathbf{1} \quad (3.16)$$

kde  $\mathbf{1}$  představuje jednotkovou matici.

Dosazením rovnice (3.15) do (3.14) získáme vztah mezi odpovídajícími si body v projektivních rovinách obou kamer. Pro možnost transformace bodů  $\mathbf{m}$  a  $\mathbf{m}'$  mezi projektivními rovinami pomocí rovnice (3.17) je nutná znalost souřadnice  $Z$  odpovídajícího bodu v 3D souřadném systému.

$$\mathbf{Z}' \tilde{\mathbf{m}}' = \mathbf{Z} \mathbf{A}' \mathbf{R} \mathbf{A}^{-1} \tilde{\mathbf{m}} + \mathbf{A}' \mathbf{t} \quad (3.17)$$



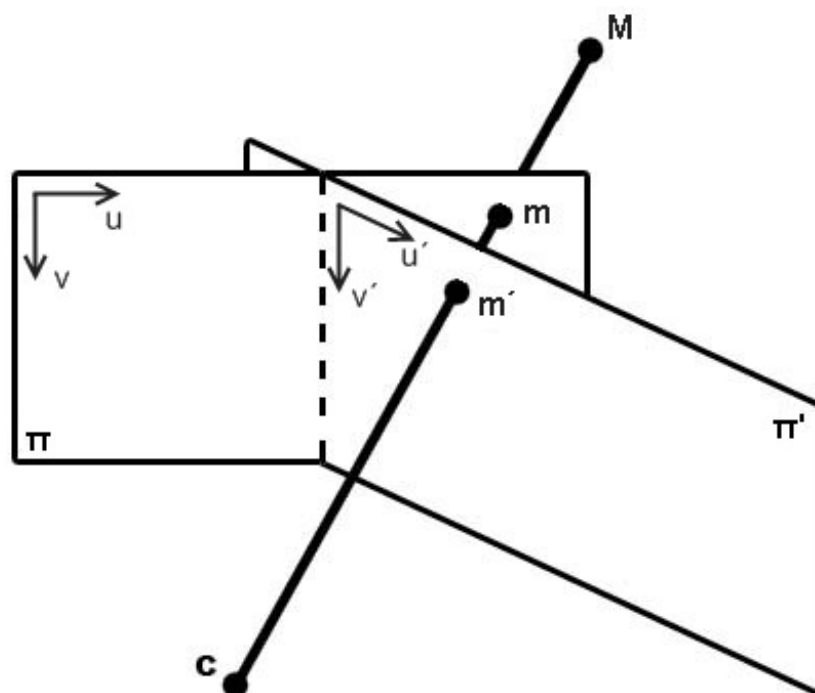
**Obr. 3.2:** Scéna zachycená dvěma kamerami



## 4. Homografie

### 4.1. Definice

Homografie je definována jako perspektivní transformace mezi dvěma projektivními rovinami  $\pi$  a  $\pi'$ . Jedná se o geometrickou transformaci popisující transformaci souřadnic bodů z roviny  $\pi$  na odpovídající souřadnice v rovině  $\pi'$ .



**Obr. 4.1:** Perspektivní transformace bodu  $M$  zachyceného v zobrazovacích rovinách  $\pi$  a  $\pi'$  perspektivní kamerou otáčenou kolem jejího optického středu ( $M$  – bod v prostoru;  $C$  – optický střed kamery;  $\pi$  – zobrazovací rovina prvního snímku;  $\pi'$  – zobrazovací rovina druhého snímku;  $m$  – bod, kde dochází k průniku roviny prvního snímku a spojnice mezi body  $X$  a  $C$ ;  $m'$  – bod, kde dochází k průniku roviny druhého snímku a spojnice mezi body  $M$  a  $C$ )

Body  $m$  a  $m'$  získané jako průniky zobrazovacích rovin  $\pi$  a  $\pi'$  a spojnice optického středu kamery  $C$  a bodu v prostoru  $M$  jsou popsány sloupcovými vektory jejich homogenních souřadnic jako

$$\mathbf{m} = [au, av, \alpha]^T, \quad (4.1)$$

$$\mathbf{m}' = [\alpha u', \alpha v', \alpha]^T, \quad (4.2)$$

kde  $\alpha$  je konstanta s podmínkou  $\alpha \neq 0$ .

Jak je patrné ze vztahů (4.1) a (4.2), nejsou u homogenních souřadnic tolik důležité číselné hodnoty, ale spíše jejich vzájemné poměry. Přepočtení homogenních souřadnic zpět na souřadnice odpovídající poloze bodu v zobrazovací rovině  $\pi$  lze vyjádřit jako

$$\mathbf{m} = \left[ \frac{\alpha u}{\alpha}, \frac{\alpha v}{\alpha}, 1 \right]^T, \quad (4.3)$$

Homografii lze popsat jako lineární transformaci homogenních souřadnic  $\mathbf{m}$  a  $\mathbf{m}'$  pomocí nesusingularní matice o rozměrech  $3 \times 3$ , tuto matici nazvěme maticí homografie  $\mathbf{H}$ . Matice homografie  $\mathbf{H}$  má 8 stupňů volnosti. Definice matice homografie  $\mathbf{H}$  vychází z rovnice (3.17). Při úpravě této rovnice se využívá faktu, že zdrojové snímky jsou pořizovány perspektivní kamerou, která se pouze otáčí kolem svého optického středu a nedochází k jejímu posunutí, díky čemuž se stává vektor  $\mathbf{t}$  nulovým. Postup při definování matice homografie  $\mathbf{H}$ :

$$\mathbf{Z}' \mathbf{m}' = \mathbf{Z} \mathbf{A}' \mathbf{R}' \mathbf{A}^{-1} \mathbf{m}, \quad (4.4)$$

$$\mathbf{m}' = \frac{\mathbf{Z}'}{\mathbf{Z}} \mathbf{A}' \mathbf{R}' \mathbf{A}^{-1} \mathbf{m}, \quad (4.5)$$

$$\mathbf{m}' = \frac{\mathbf{Z}'}{\mathbf{Z}} \mathbf{H} \mathbf{m}, \quad (4.6)$$

$$\alpha \mathbf{m}' = \mathbf{H} \mathbf{m}, \quad (4.7)$$

$$\mathbf{m}' \simeq \mathbf{H} \mathbf{m}, \quad (4.8)$$

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \simeq \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \quad (4.9)$$

## 4.2. Výpočet homografie

Matici homografie  $\mathbf{H}$  lze vypočítat za pomoci bodových nebo přímkových mezisnímkových korespondencí. Obě metody jsou ekvivalentní. V praxi se setkáme spíše s výpočtem homografie pomocí bodové korespondence, protože i algoritmy hledání mezisnímkových korespondencí pracují s významnými body.

### 4.2.1. Bodová korespondence

Všechny zjištěné mezisnímkové korespondence si označíme indexem  $i = 1, \dots, n$ , kde  $n$  je počet párů. Potom lze rovnici (4.7) přepsat do tvaru

$$\alpha_i \mathbf{m}'_i = \mathbf{H} \mathbf{m}_i . \quad (4.10)$$

a následně rozepsat jako

$$\alpha_i u'_i = \mathbf{h}_1^T \mathbf{m}_i , \quad (4.11)$$

$$\alpha_i v'_i = \mathbf{h}_2^T \mathbf{m}_i , \quad (4.12)$$

$$\alpha_i = \mathbf{h}_3^T \mathbf{m}_i , \quad (4.13)$$

kde  $\mathbf{h}_1^T$ ,  $\mathbf{h}_2^T$  a  $\mathbf{h}_3^T$  jsou jednotlivé řádky matice homografie  $\mathbf{H}$ .

Dosazením rovnice (4.13) do rovnic (4.11) a (4.12) získáme soustavu dvou lineárních algebraických rovnic

$$\mathbf{h}_3^T \mathbf{m}_i u'_i = \mathbf{h}_1^T \mathbf{m}_i , \quad (4.14)$$

$$\mathbf{h}_3^T \mathbf{m}_i v'_i = \mathbf{h}_2^T \mathbf{m}_i . \quad (4.15)$$

Spojením rovnic (4.14) a (4.15) pro všechny korespondující body získáme systém  $2n$  homogenních lineárních algebraických rovnic, který po přepsání do maticového tvaru vypadá

$$\begin{bmatrix} \mathbf{m}_i^T & 0 & 0 & 0 & -u'_i \mathbf{m}_i^T \\ 0 & 0 & 0 & \mathbf{m}_i^T & -v'_i \mathbf{m}_i^T \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix} = 0, \quad (4.16)$$

$$\mathbf{A}\mathbf{h}=0. \quad (4.17)$$

Pro výpočet matice homografie  $\mathbf{H}$  je nutné, aby počet lineárně nezávislých řádků matice  $\mathbf{A}$  byl minimálně stejný jako počet stupňů volnosti matice homografie  $\mathbf{H}$ . Protože počet řádků matice  $\mathbf{A}$  je definován jako  $2n$ , kde  $n$  je počet korespondujících bodů, pak jsou potřeba minimálně čtyři korespondující body, z nichž žádné tři body nesmí ležet na přímce, což odpovídá podmínce lineární nezávislosti jednotlivých řádků matice  $\mathbf{A}$ .

Řešením soustavy (4.17) je nulový prostor matice  $\mathbf{A}$ , čehož dosáhneme SVD rozkladem této matice, viz [4, str. 159]. Chceme-li přesnější a odolnější výpočet je vhodné volit  $n \gg 4$ .

#### 4.2.2. Přímková korespondence

Matice homografie  $\mathbf{H}$  lze také určit ze znalosti vektorů korespondujících přímek zdrojových snímků. Korespondující přímky se transformují podle

$$\mathbf{H}^T \mathbf{l}' = \beta \mathbf{l}, \quad (4.18)$$

kde  $\beta$  je konstanta s podmínkou  $\beta \neq 0$ ,  $\mathbf{l}$  a  $\mathbf{l}'$  jsou homogenní reprezentace vektorů korespondujících přímek zdrojových snímků definovaných jako

$$\mathbf{l} = [\beta u, \beta v, \beta w]^T, \quad (4.19)$$

$$\mathbf{l}' = [\beta u', \beta v', \beta w']^T. \quad (4.20)$$

Při volbě  $\beta = 1$  odpovídá homogenní reprezentace přímky reprezentaci přímky v zobrazovací rovině  $\pi$ .

Všechny páry korespondujících přímek označíme indexem  $i=1, \dots, n$ , kde  $n$  je počet párů. Potom lze dle (4.18) psát

$$\mathbf{H}^T \mathbf{l}'_i = \beta_i \mathbf{l}_i \quad . \quad (4.21)$$

Rovnici (4.21) lze rozepsat jako

$$\mathbf{g}_1^T \mathbf{l}'_i = \beta_i u_i \quad , \quad (4.22)$$

$$\mathbf{g}_2^T \mathbf{l}'_i = \beta_i v_i \quad , \quad (4.23)$$

$$\mathbf{g}_3^T \mathbf{l}'_i = \beta_i w_i \quad , \quad (4.24)$$

kde  $\mathbf{g}_1$ ,  $\mathbf{g}_2$  a  $\mathbf{g}_3$  jsou jednotlivé sloupce matice homografie  $\mathbf{H}$ .

Abychom z předchozích rovnic odstranili neznámou konstantu  $\beta_i$  podělíme mezi sebou vždy dvě ze tří rovnic (4.22), (4.23) a (4.24), čímž dostaneme

$$\frac{\mathbf{g}_1^T \mathbf{l}'_i}{\mathbf{g}_3^T \mathbf{l}'_i} = \frac{\beta_i u_i}{\beta_i w_i} \quad , \quad (4.25)$$

$$\frac{\mathbf{g}_2^T \mathbf{l}'_i}{\mathbf{g}_3^T \mathbf{l}'_i} = \frac{\beta_i v_i}{\beta_i w_i} \quad , \quad (4.26)$$

$$\frac{\mathbf{g}_1^T \mathbf{l}'_i}{\mathbf{g}_2^T \mathbf{l}'_i} = \frac{\beta_i u_i}{\beta_i v_i} \quad . \quad (4.27)$$

Rovnice (4.25), (4.26) a (4.27) lze jednoduše upravit do tvaru soustavy tří homogenních lineárně algebraických rovnic pro vektory sloupců matice  $\mathbf{H}$

$$\mathbf{l}'_i \mathbf{g}_1^T w_i - \mathbf{l}'_i \mathbf{g}_3^T u_i = 0 \quad , \quad (4.28)$$

$$\mathbf{l}'_i \mathbf{g}_2^T w_i - \mathbf{l}'_i \mathbf{g}_3^T v_i = 0 \quad , \quad (4.29)$$

$$\mathbf{l}'_i \mathbf{g}_1^T v_i - \mathbf{l}'_i \mathbf{g}_2^T u_i = 0 \quad . \quad (4.30)$$

Spojením rovnic (4.28), (4.29) a (4.30) pro všechny korespondující přímky získáme systém  $3n$  homogenních lineárních algebraických rovnic, kde  $n$  je počet korespondujících párek, který po přepsání do maticového tvaru vypadá

$$\begin{bmatrix} w_i \mathbf{l}_i'^T & 0 & 0 & 0 & -u_i \mathbf{l}_i'^T \\ 0 & 0 & 0 & w_i \mathbf{l}_i'^T & -v_i \mathbf{l}_i'^T \\ v_i \mathbf{l}_i'^T & -u_i \mathbf{l}_i'^T & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \mathbf{g}_3 \end{bmatrix} = 0 \quad , \quad (4.31)$$

$$\mathbf{B}\mathbf{g}=0 \quad . \quad (4.32)$$

Pro výpočet matice homografie  $\mathbf{H}$  je nutné, aby počet lineárně nezávislých řádků matice  $\mathbf{B}$  byl minimálně stejný jako počet stupňů volnosti matice homografie  $\mathbf{H}$ . Protože počet řádků matice  $\mathbf{B}$  je definován jako  $3n$ , kde  $n$  je počet korespondujících přímek, pak jsou potřeba minimálně tři vektory korespondujících přímek. Žádné dvě přímky nesmí mít stejnou polohu v zobrazovací rovině  $\pi$ ., jinak by nebyla splněna podmínka lineární nezávislosti řádků matice  $\mathbf{B}$ .

Řešením soustavy (4.32) je nulový prostor matice  $\mathbf{B}$ , čehož dosáhneme SVD rozkladem této matice, viz [4, str. 159]. Chceme-li přesnější a odolnější výpočet je vhodné volit  $n \gg 3$ .

## 5. Skládání zdrojových snímků

Známe-li matici homografie  $H$  můžeme přejít k dalšímu kroku, čímž je transformace zdrojových snímků a jejich následné spojení do panoramatického obrazu.

Na obr. 5.1 máme zdrojové snímky, které za pomoci vytvořené aplikace spojíme do panoramatického pohledu. Jen pro úplnost jsou na obr. 5.2 uvedeny zdrojové snímky po procesu transformace do společné zobrazovací roviny.



**Obr. 5.1:** Zdrojové snímky



**Obr. 5.2:** Transformované snímky před spojením do panoramatického obrazu



**Obr. 5.3:** Výsledné panorama sestavené přeložením jednoho snímku přes druhý

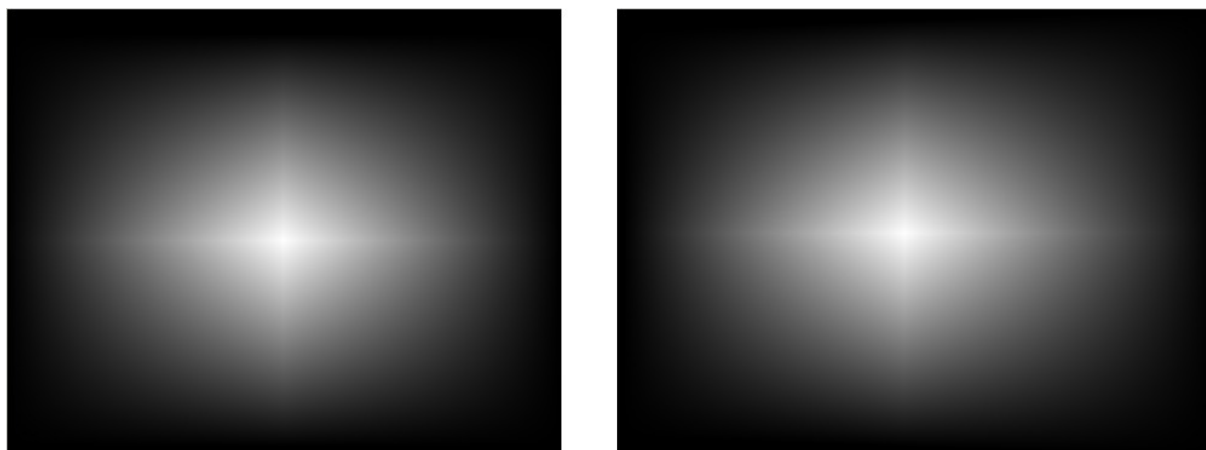
Na obr. 5.3 je vidět nežádoucí ostrý přechod mezi zdrojovými snímky ve výsledném panoramatickém obraze, který vznikne pouhým překrytím transformovaných zdrojových snímků přes sebe. Tento ostrý přechod vzniká během transformace při procesu interpolace, protože obrazové body zdrojového snímku nejsou transformovány přesně do rastru obrazových bodů cílového snímku. Toto lze odstranit některou z níže popsaných metod.

Pro metody ostraňování nežádoucích ostrých přechodů mezi zdrojovými snímky ve výsledném panoramatickém obraze jsou nejdůležitější matice vah  $\mathbf{W}$  a  $\mathbf{W}_{\text{MAX}}$ , které jsou stejných rozměrů jako odpovídající zdrojový snímek, kterému tyto matice náleží.

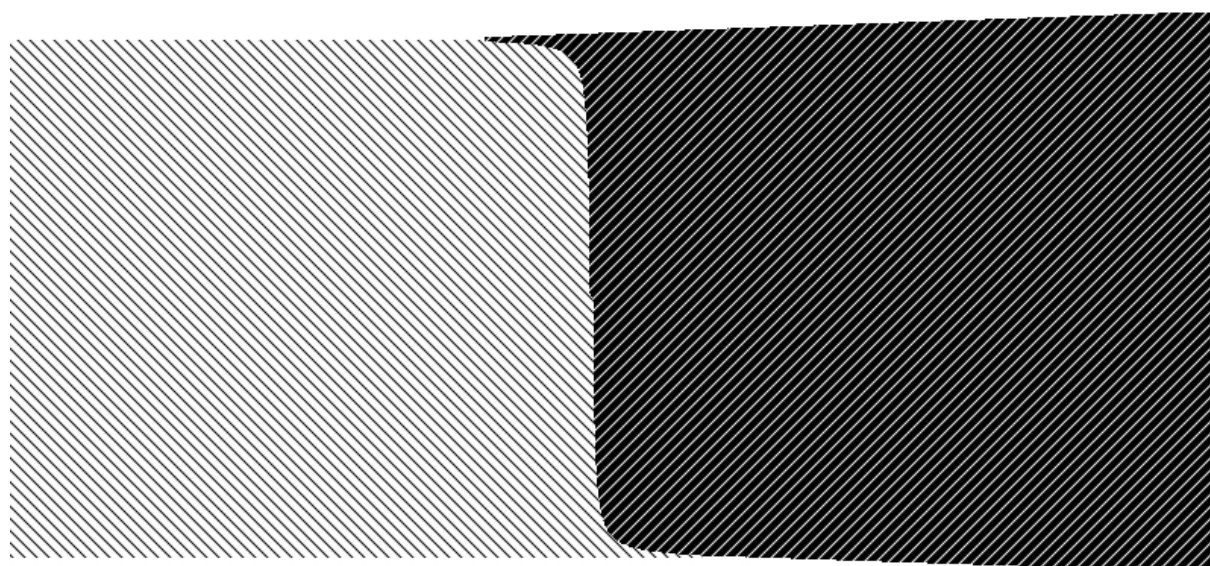
$$\mathbf{W}(x, y) = \mathbf{v}(x) \mathbf{w}(y) \quad , \quad (5.1)$$

kde  $\mathbf{v}$  a  $\mathbf{w}$  jsou vektory, jejichž hodnota lineárně roste od nuly po jedna ve středu vektoru a pak zpět klesá k nule.  $\mathbf{W}_{\text{MAX}}(x, y)$  je rovna 1, pokud  $\mathbf{W}(x, y)$  je větší než  $\mathbf{W}(u, v)$  všech ostatních zdrojových snímků, jinak je rovna 0. Souřadnice  $(x, y)$  a  $(u, v)$  sice určují polohu ve více různých snímcích, ale popisují polohu stejného bodu zobrazované scény.





**Obr. 5.4:** Váhová matice  $W$  pro zdrojové snímky z obr. 5.1



**Obr. 5.5:** Váhová matice  $W_{MAX}$  vycházející z váhových matic  $W$  dle obr. 5.4

Nejjednodušší metodou jak složit výsledný panoramatický obraz je vymaskování transformovaných zdrojových snímků (obr. 5.2) a jejich následné složení dle obr. 5.5. Výsledek je vidět na obr. 5.6. Nežádoucí ostrý přechod mezi zdrojovými snímky zde již sice není tolik patrný, ale stále není úplně minimalizován. Pokročilejší metody využívají principu prolínání zdrojových snímků. V závislosti na volbě parametrů prolínání zdrojových snímků se liší i výsledné panoramatické snímky.



**Obr. 5.6:** Výsledný panoramatický snímek dle obr. 5.5

### **5.1. Lineární prolínání**

U metody lineárního prolínání můžeme pro oblast, kde se zdrojové snímky překrývají psát

$$\mathbf{I}(u, v) = \frac{\sum_{i=1}^n \mathbf{W}_i(x_i, y_i) \mathbf{I}_i(x_i, y_i)}{\sum_{i=1}^n \mathbf{W}_i(x_i, y_i)}, \quad (5.2)$$

kde  $n$  je počet snímků, které se vzájemně překrývají ve výsledném panoramatickém snímku,  $\mathbf{I}$  značí jasovou matici odpovídajícího snímku a  $\mathbf{W}$  jeho váhovou matici dle (5.1),  $(u, v)$  jsou souřadnice obrazového bodu ve výsledném panoramatickém obraze,  $(x, y)$  souřadnice obrazového bodu zdrojového snímku. I když všechny souřadnice popisují polohu obrazového bodu v jiném zdrojovém snímku, stále ukazují na tentýž bod zobrazované scény.



**Obr. 5.7:** Lineární prolínání transformovaných zdrojových snímků z obr. 5.2



**Obr. 5.8:** Výsledné panoramatický snímek s použitím lineárního prolnutí

## ***5.2. Multipásmové prolínání***

Představme si zdrojové snímky jako kmitočtová spektra. Tato metoda spočívá v rozdělení těchto spekter na několik dílčích kmitočtových pásem. Odpovídající si kmitočtová pásma jednotlivých zdrojových snímků jsou poté skládána podle (5.10), čímž jsou získávána dílčí kmitočtová pásma výsledného panoramatického obrazu. Máme-li zpracovány všechna kmitočtová pásma, nezbyváá než tyto dílčí kmitočtová pásma sečíst, čímž získáme výsledný panoramatický snímek.

Hlavní výhoda této metody je využití rozdílné šířky oblasti prolinání pro každé kmitočtové pásmo. Pro vysoké kmitočty je oblast prolinání úzká, kdežto pro nízké kmitočty naopak široká.

Pro oddělení jednotlivých kmitočtových pásem se používá konvoluce jasové matice s Gaussovou funkcí, která se chová jako dolní propust.

Jasovou matici pro složky s nejvyššími kmitočty obrazu lze popsat jako

$$\mathbf{B}_\sigma(x, y) = \mathbf{I}(x, y) - \mathbf{I}_\sigma(x, y) \quad , \quad (5.3)$$

$$\mathbf{I}_\sigma(x, y) = \mathbf{I}(x, y) * \mathbf{G}(x, y, \sigma) \quad , \quad (5.4)$$

kde  $\mathbf{G}$  je Gaussova funkce definovaná v (2.12), symbol  $*$  značí konvoluci a  $\mathbf{B}_\sigma$  představuje jasovou matici obrazových složek s kmitočtem vyjádřených pomocí vlnové délky  $\lambda \in [0; \sigma]$ .

Dále si definujeme matici vah  $\mathbf{W}_\sigma(x, y)$  pro stejné kmitočtové pásmo jako

$$\mathbf{W}_\sigma(x, y) = \mathbf{W}_{\text{MAX}}(x, y) * \mathbf{G}(x, y, \sigma) \quad . \quad (5.5)$$

Rovnice (5.3), (5.4) a (5.5) si vyjádříme pro další kmitočtová pásma jako

$$\mathbf{B}_{(k+1)\sigma}(x, y) = \mathbf{I}_{k\sigma}(x, y) - \mathbf{I}_{(k+1)\sigma}(x, y) \quad , \quad (5.6)$$

$$\mathbf{I}_{(k+1)\sigma}(x, y) = \mathbf{I}_{k\sigma}(x, y) * \mathbf{G}(x, y, \sigma') \quad , \quad (5.7)$$

$$\mathbf{W}_{(k+1)\sigma}(x, y) = \mathbf{W}_\sigma(x, y) * \mathbf{G}(x, y, \sigma') \quad , \quad (5.8)$$

$$\sigma' = \sigma \sqrt{2k+1} \quad , \quad (5.9)$$

kde  $k$  je konstanta vyjadřující kmitočtové pásmo, pro kterou platí  $k \geq 1$ .

Nyní si můžeme vyjádřit jasovou matici  $\mathbf{I}_{k\sigma}$  pro každé kmitočtové pásmo  $k$

$$\mathbf{I}_{k\sigma}(u, v) = \frac{\sum_{i=1}^n \mathbf{B}_{k\sigma}^i(x_i, y_i) \mathbf{W}_{k\sigma}^i(x_i, y_i)}{\sum_{i=1}^n \mathbf{W}_{k\sigma}^i(x_i, y_i)}, \quad (5.10)$$

kde  $n$  je počet snímků, které se vzájemně překrývají ve výsledném panoramatickém snímku,  $\mathbf{B}_{k\sigma}$  značí jasovou matici odpovídajícího kmitočtového pásma zdrojového snímku a  $\mathbf{W}_{k\sigma}$  jeho váhovou matici dle (5.5) nebo (5.8),  $(u, v)$  jsou souřadnice obrazového bodu ve výsledném panoramatickém obraze,  $(x, y)$  souřadnice obrazového bodu zdrojového snímku. Ikdyž všechny souřadnice popisují polohu obrazového bodu v jiném zdrojovém snímku, stále ukazují na tentýž bod zobrazované scény, index  $i$  vyjadřuje pořadové číslo zdrojového snímku, nikoliv  $i$ -tou mocninu.

Jasová matice výsledného panoramatického obrazu vznikne sečtením jasových matic všech kmitočtových pásem dle

$$\mathbf{I}(u, v) = \sum_{\forall k} \mathbf{I}_{k\sigma}(u, v) \quad (5.11)$$

Obr. 5.9 – 5.12 popisují skládání panoramatického obrazu při užití metody multipásmové prolínání zdrojových snímků. Vlevo na obr. 5.10 – 5.12 jsou zobrazeny jasové matice dílčích kmitočtových pásem zdrojových snímků  $\mathbf{B}_{k\sigma}$  a vpravo jsou zobrazeny jejich váhové matice  $\mathbf{W}_{k\sigma}$ , kde černá značí nulovou váhu a bílá značí váhu maximální. Porovnáme-li pravé části obr. 5.10 – 5.12 zjistíme, že s rostoucí velikostí konstanty  $k$  roste i velikost oblasti prolínání jednotlivých zdrojových snímků.



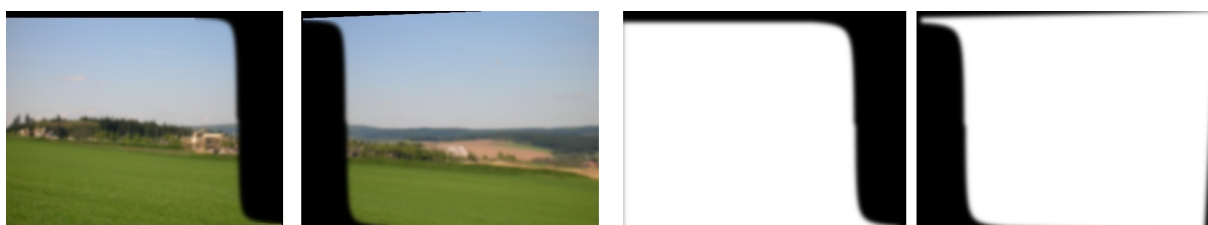
**Obr. 5.9:** Zdrojové snímky a výsledné panorama



**Obr. 5.10:** Multipásmové prolínání,  $k = 1$



**Obr. 5.11:** Multipásmové prolínání,  $k = 2$



**Obr. 5.12:** Multipásmové prolínání,  $k = 3$

### **5.3. Porovnání metod**

Na následujících čtyřech obrázcích můžeme vidět různé panoramatické snímky spojené pomocí obou metod. Vrchní snímek zachycuje panoramatický snímek spojený pomocí metody lineárního prolínání a spodní pomocí metody multipásmového prolínání.

Na obr. 5.13, 5.15 a 5.16 je vidět, že výsledné panoramatické snímky spojené odlišnými metodami se od sebe neliší. Na obr. 5.14 jsou již odlišnosti mezi metodami patrné. Tyto odlišnosti jsou způsobeny jasovou odlišností zdrojových snímků. Metoda lineárního prolínání se vyznačuje prolínáním přes celou oblast překryvu zdrojových snímků ve výsledném panoramatickém obraze (viz obr. 5.7). Naopak u metody multipásmového prolínání je oblast překryvu zdrojových snímků ve výsledném panoramatickém obraze podstatně užší (viz obr. 5.10 – 5.12). Právě vlivem šířky oblasti prolínání vychází lépe metoda lineárního prolínání, respektive u metody lineárního prolínání je přechod mezi zdrojovými snímky ve výsledném panoramatickém obraze méně patrný.





**Obr. 5.13:** Panoramatický snímek 1 (5 zdrojových snímků)



**Obr. 5.14:** Panoramatický snímek 2 (3 zdrojové snímky)



**Obr. 5.15:** Panoramatický snímek 3 (3 zdrojové snímky)

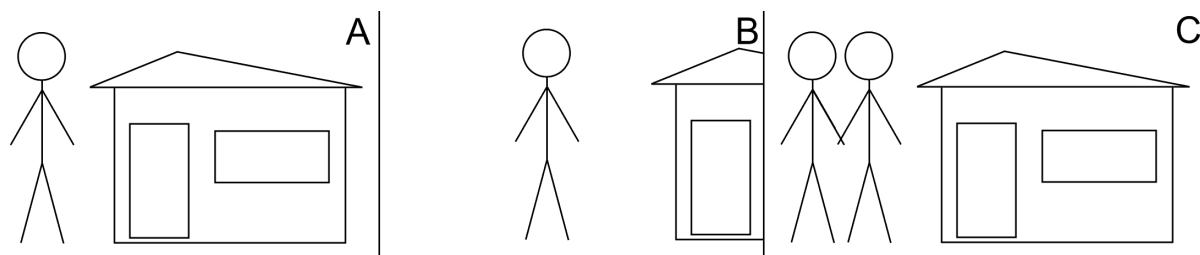


**Obr. 5.16:** Panoramatický snímek 4 (2 zdrojové snímky)



## 5.4. Chyba paralaxe

Pokud se liší poloha optického středu a poloha středu otáčení kamery při pořizování zdrojových snímků, mohou se lišit vzájemné polohy objektů popředí a pozadí scény zachycených kamerou při různých úhlech natočení kamery.



Obr. 5.17: Chyba paralaxe

Na obr. 5.17A je zachycena postava jako objekt v popředí a budova jako objekt v pozadí scény. Vlivem rozdílné polohy optického středu a středu otáčení kamery se na obr. 5.17B může zdát, že došlo ke změně vzájemné polohy postavy a budovy, jedná se ale o stejnou scénu jako na obr. 5.17A, ovšem tato scéna je zachycena kamerou, které byla otočena o úhel  $\theta$ . Obr. 5.17C zachycuje snímek složený ze zdrojových snímků (obr. 5.17A a obr. 5.17B). Vlivem různých vzájemných poloh postavy a budovy ve zdrojových snímcích, bude postava ve výsledném snímku zobrazena dvojitě, což nelze odstranit jinak než pořízením nových zdrojových snímků, kdy bude kladen větší důraz na minimalizaci vzdálenosti optického středu a středu otáčení kamery. V praxi se k minimalizaci této vzdálenosti používají panoramatické hlavy, které se připevňují mezi kameru a stativ.

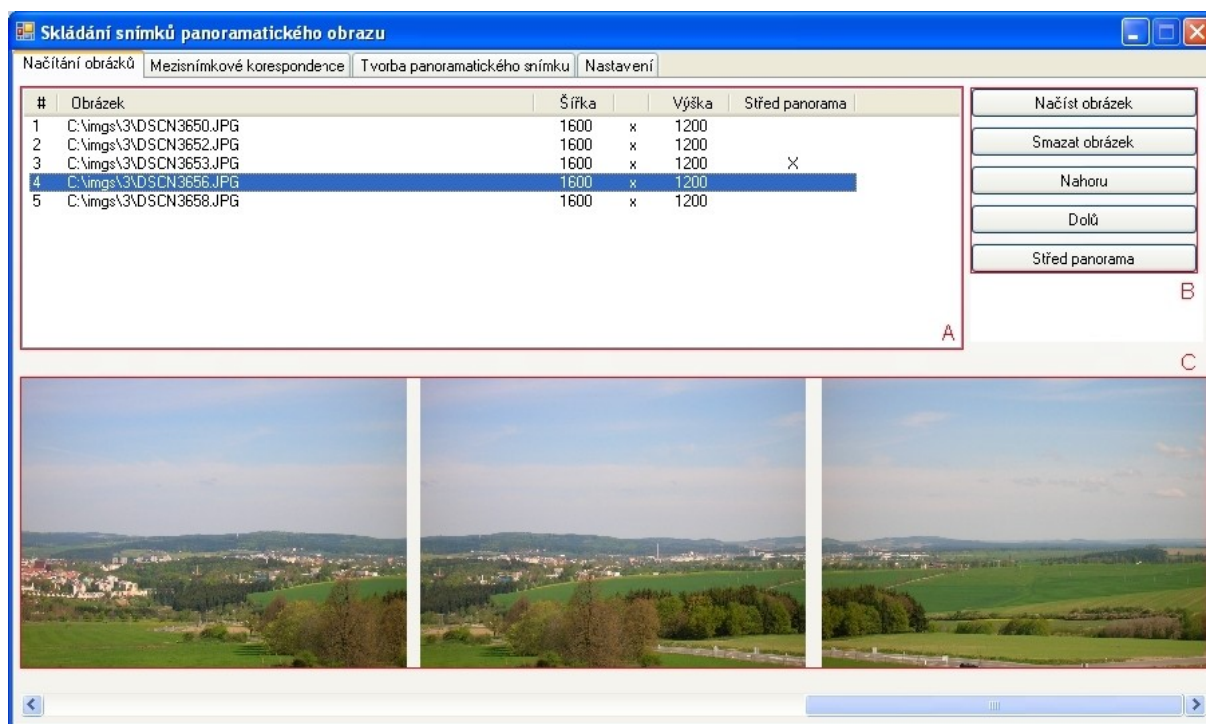
Chyba paralaxe se také neprojeví, pokud je vzdálenost mezi optickým středem a středem otáčení kamery vůči vzdálenosti kamery a zachycených objektů scény zanedbatelná.

## 6. Vytvořená aplikace

Aplikace je vytvořena pomocí doporučených nástrojů, kterými jsou Visual C++ 2005 Express Edition a knihovna OpenCV pro zpracování obrazu. Jedná se o aplikaci určenou pro operační systémy MS Windows.

### 6.1. Načítání snímků

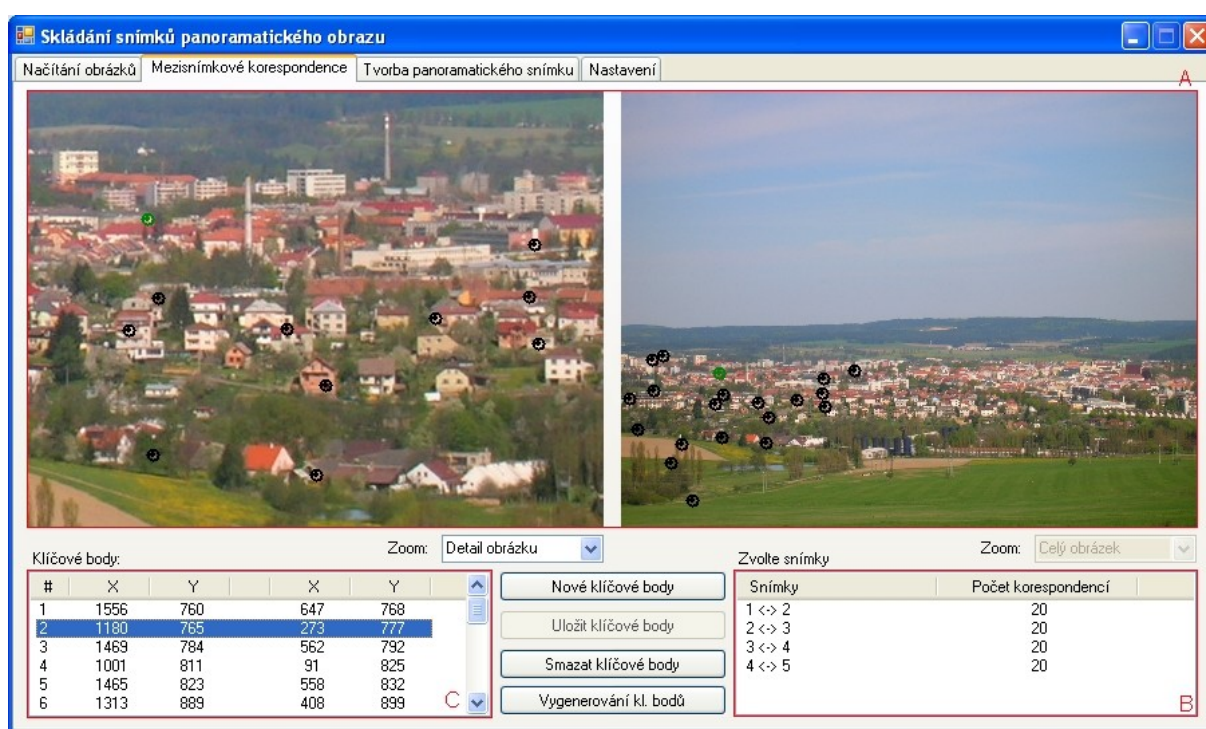
Aplikace se ovládá přes jednoduché grafické rozhraní (obr. 6.1). Okno aplikace při načítání zdrojových snímků je rozděleno na tři hlavní části. Informace o načtených snímcích (obr. 6.1, A), funkční tlačítka (obr. 6.1, B) a náhled načtených snímků (obr. 6.1, C). Během této fáze se také volí snímek, který bude brán ve fázi tvorby panoramatického obrazu jako jeho střed. Pro správnou funkci aplikace se předpokládá, že zdrojové snímky budou po načtení správně seřazeny. Správnost posloupnosti zdrojových snímků si lze zkontrolovat ve spodní části okna (obr. 6.1, C).



Obr. 6.1: Načítání zdrojových snímků

## 6.2. Mezisnímkové korespondence

Aplikace umožňuje vytvářet mezisnímkové korespondence dvěma způsoby. První možností je ruční vytváření pomocí levého tlačítka myši přímo do náhledu zdrojových snímků (obr. 6.2, A). Aplikace také umožňuje zobrazení detailního náhledu zdrojového snímku, ve kterém se lze pohybovat pomocí pravého tlačítka myši. Okno aplikace obsahuje informace o počtu mezisnímkových korespondencí pro všechny páry zdrojových snímků (obr. 6.2, B) a detailní výpis poloh korespondujících bodů pro zvolený pár zdrojových snímků (obr. 6.2, C).



Obr. 6.2: Mezisnímkové korespondence

Druhou možností pro získání mezisnímkových korespondencí představuje využití implementovaného algoritmu SIFT. Automatické vygenerování mezisnímkových korespondencí je možné spustit kliknutím na tlačítko „Vygenerování kl. bodů“. Aplikace využívá implementaci SIFT algoritmu vytvořenou Robem Hessem [10]. Rob Hess implementoval SIFT algoritmus v jazyce C, proto bylo nutné zdrojové kódy upravit kvůli drobným odlišnostem syntaxí programovacích jazyků C a C++. Implementace SIFT algoritmu od Roba Hesse je šířena pod licencí GNU GPL [11].

V záložce „Nastavení“ je možné měnit klíčové parametry automatického získávání mezisnímkových korespondencí. Jedná se o práh pro rozhodování zda jsou dva významné body různých zdrojových snímků korespondující. Druhým parametrem důležitým pro SIFT je velikost zdrojového obrazu, ve kterém se hledají významné body. Protože se metoda SIFT vyznačuje odolností vůči změně měřítka, není nutné hledat významné body ve zdrojových snímcích v plném rozlišení, ale je možné rozlišení zdrojového snímku snížit, s čímž souvisí i snížení výpočetní náročnosti. Další možností v nastavení týkající se procesu získávání mezisnímkových korespondencí je volba zapnutí či vypnutí filtrace RANSAC a velikost její tolerance. Poslední volbou je maximální počet mezisnímkových korespondencí, které budou uloženy a se kterými se bude dále počítat.

### **6.3. Tvorba panoramatického snímku**

Základním snímkem při tvorbě panoramatického snímku je pravý okrajový zdrojový snímek. K tomuto snímku se postupně připojují všechny ostatní zdrojové snímky.

Proces tvorby výsledného panoramatického snímku je tvořen třemi základními kroky, které si blíže popíšeme. Máme-li zjištěn minimální počet správných mezisnímkových korespondencí, můžeme přejít k prvnímu kroku, kterým je výpočet matice homografie  $\mathbf{H}$ . V aplikaci je implementován výpočet homografie pomocí bodových korespondencí. Samotný výpočet využívá zapouzdřené metody `cvFindHomography` z knihoven OpenCV. Této metodě jsou předávány souřadnice korespondujících bodů zdrojových snímků, mezi kterými se homografie počítá, a ukazatel na matici o rozměrech  $3 \times 3$ , do které je matice homografie  $\mathbf{H}$  uložena.

Dalším krokem je výpočet rozměrů výsledného panoramatického snímku. Tyto rozměry se počítají ze souřadnic rohů dílčích zdrojových snímků, protože rohy představují mezní body zdrojových snímků. Při výpočtu velikosti výsledného panoramatického snímku se provádí testovací transformace každého z rohů právě skládaných zdrojových snímků pomocí odpovídající transformační matice a zkoumá se, na jaké souřadnice v rovině výsledného panoramatického snímku se daný roh promítne. V některých případech se totiž při transformaci zdrojových snímků může stát, že by byly obrazové informace transformovány do

roviny panoramatického snímku na pozici se zápornými souřadnicemi, což je nežádoucí. Proto je nutné si definovat matice posunutí  $\mathbf{T}$  jako

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (6.1)$$

kde  $t_x$  představuje posunutí v ose x a  $t_y$  posunutí v ose y v rovině panoramatického snímku. Dodatečnou transformací obrazových dat zdrojových snímků touto maticí lze obrazová data přesunout do nezáporných souřadnic polohy ve výsledném panoramatickém snímku. Abychom ušetřili výpočetní výkon je vhodné sloučit obě transformace v jednu, a to pomocí maticového součinu dle

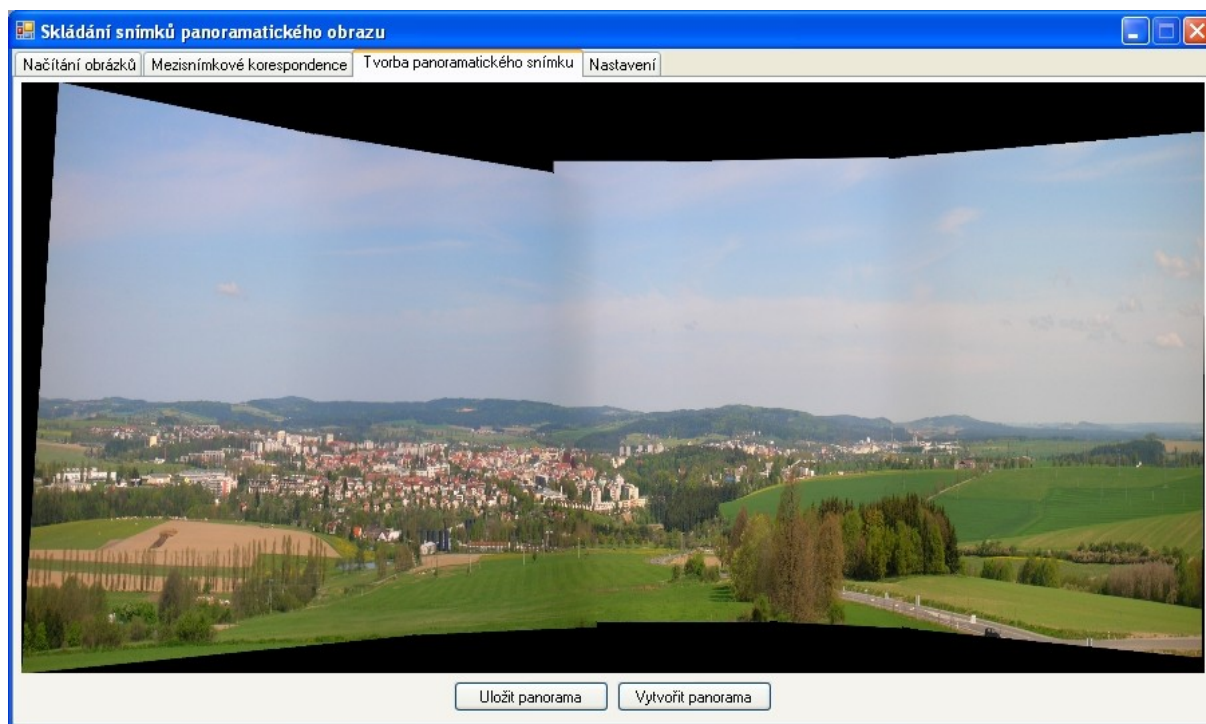
$$\mathbf{H}' = \mathbf{T} \mathbf{H}. \quad (6.2)$$

Aby se při výpočtu velikosti výsledného panoramatického snímku zohlednily i rozměry zdrojových snímků, které již byli připojeny, jsou souřadnice rohů všech složených zdrojových snímků uloženy do dynamického seznamu v datové struktuře typu vector. Nejprve jsou do tohoto seznamu uloženy souřadnice rohů základního snímku, ke kterým se postupně přidávají souřadnice rohů snímku, který byl zrovna připojen do panoramatického obrazu. V každém kroku připojení dalšího zdrojového snímku k mezivýslednému panoramatickému snímku jsou uloženy souřadnice rohů zdrojových snímků transformovány stejně jako samotný mezivýsledný panoramatický snímek.

Následuje transformace zdrojových snímků pomocí matice homografie  $\mathbf{H}'$ , dle (6.2). Aplikace při transformaci obrazových dat zdrojových snímků využívá zapouzdřené metody `cvWarpPerspective` z knihoven OpenCV. Této metodě se jako parametry předají ukazatel na obrazová data snímku, který chceme transformovat, a ukazatel na obrazová data snímku, do kterého se má transformovaný snímek uložit. Posledním předávaným parametrem je samotná matice homografie  $\mathbf{H}'$ .

Posledním krokem je spojení transformovaných zdrojových snímků. V aplikaci jsou implementovány obě popsané metody odstranění nežádoucích ostrých přechodů mezi

zdrojovými snímky ve výsledném panoramatickém obraze. Volbu metody, jenž bude použita při skládání transformovaných snímků, lze provést v záložce „Nastavení“. U metody multipásmového prolínání je potřeba zmínit, že se vytváří pro každé pásmo samostatný panoramatický snímek. Tyto dílčí panoramatické snímky jsou nakonec sloučeny dle (5.11) v jeden výsledný panoramatický snímek, který je zobrazen v okně aplikace (obr. 6.3). V spodní části okna aplikace se nachází i dvě tlačítka. První slouží k uložení vytvořeného panoramatického snímku a druhé k opakování procesu vytvoření panoramatického snímku. K opakování procesu vytvoření panoramatického snímku dochází také vždy při zobrazení záložky „Tvorba panoramatického snímku“, pokud byl změněn některý z parametrů majících na tento proces vliv.



Obr. 6.3: Tvorba panoramatického snímku

## 6.4. Zdrojové kódy

Deklarace vlastních tříd IMGS, IMG a IMG\_POINT jsou uloženy v souborech „imgs.h“ a „imgs.cpp“. Třída IMGS slouží k uložení informací o zdrojových snímcích a důležitých parametrech výsledného panoramatického snímku. Třída IMG slouží k uložení informací o každém zdrojovém snímku a o korespondencích s ostatními zdrojovými snímky. Poslední třída IMG\_POINT slouží k uložení souřadnic klíčových bodů.

Zdrojové kódy grafického rozhraní a ovládání aplikace, stejně jako veškeré metody potřebné při procesu skládání snímků panoramatického pohledu, jsou uloženy v souboru „Form1.h“.

V ostatních souborech jsou uloženy datové struktury a metody převzaté implementace algoritmu SIFT od Roba Hesse. Jedná se o soubory „imgfeatures.h“, „imgfeatures.cpp“, „kdtree.h“, „kdtree.cpp“, „minpq.h“, „minpq.cpp“, „sift.h“, „sift.cpp“, „utils.h“ a „utils.cpp“.

## 7. Závěr

V této práci jsou podrobně popsány všechny algoritmy potřebné při procesu skládání panoramatického obrazu z několika dílčích snímků scény pořízených perspektivní kamerou otáčenou kolem jejího optického středu. Základem pro úspěšné složení panoramatického obrazu je získání správných a hlavně přesných mezisnímkových korespondencí. I přes drobné nepřenosti je sice možné panoramatický obraz úspěšně složit, ale zdrojové snímky na sebe nebudou správně napojeny. Kvůli požadované přesnosti jsou používány algoritmy automatického generování mezisnímkových korespondencí. Historicky starší detektory významných bodů, jako je Moravcův a Harrisův detektor, se pro své nevýhody již moc nevyužívají, daleko častěji se setkáme s detektory SIFT, které nabízejí měřítkovou nezávislost a vysokou stabilitu nalezených významných bodů.

Na základě testování metod odstraňování nežádoucích ostrých přechodů mezi zdrojovými snímky ve výsledném panoramatickém obraze v kapitole 5.3., kdy nebylo prokázáno, která z metod nabízí lepší výsledky, lze metody lineárního a multipásmového prolínání označit jako ekvivalentní. I přes ekvivalenci výsledků obou metod doporučuji spíše metodu lineárního prolínání. Důvody jsou dva. Budou-li zdrojové snímky méně kvalitní (viz obr. 5.14), pak díky větší šířce oblasti prolínání zdrojových snímků bude ostrý přechod mezi zdrojovými snímky pozvolnější a tím i méně patrný. Druhým důvodem je výpočetní náročnost. Výpočetní náročnost metody lineárního prolínání je  $n$ -krát nižší než metoda multipásmového prolínání, kde  $n$  představuje počet kmitočtových pásem, na které jsou zdrojové snímky rozloženy.

Dalším bodem zadání je vytvoření samostatné aplikace implementující získané teoretické znalosti. Jak již bylo řečeno v textu, jedná se o aplikaci určenou pro operační systémy typu MS Windows a napsanou v jazyce C++ s využitím knihoven OpenCV. Navíc je v aplikaci využita převzatá implementace SIFT algoritmu od Roba Hesse.

Počet snímků, které dokáže vytvořená aplikace zpracovat je omezen pouze velikostí operační paměti počítače. Z tohoto důvodu jsem při vytváření aplikace kladl důraz hlavně na její nízké paměťové nároky, čímž se zvýší počet zdrojových snímků, jenž je možné



spojit. Nízkých paměťových nároků lze dosáhnout pouze na úkor rychlosti procesu skládání výsledného panoramatického snímku. Úspora operační paměti spočívá v tom, že jsou v aplikaci načteny pouze informace o rozměrech zdrojových snímků a o fyzickém umístění obrazových dat na datovém nosiči. K samotnému načtení obrazových dat zdrojových snímků dochází až pokud jsou tyto data zpracovávána. Zvýšené nároky na výpočetní výkon potřebný při tvorbě panoramatického snímku jsou vzhledem k výpočetním kapacitám dnešních počítačů zanedbatelné.

## 8. Použitá literatura

- [1] GONZALES, Rafael, WOODS, Richard. *Digital Image Processing*. [s.l.] : [s.n.], 2002. 793 s.
- [2] FISHER, Robert. *CVonline: The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision* [online]. c2007 [cit. 2007-12-18]. Dostupný z WWW: <<http://homepages.inf.ed.ac.uk/rbf/CVonline/>>.
- [3] PRATA, Stephen. *Mistrovství v C++*. [s.l.] : [s.n.], 2004. 1005 s.
- [4] DOSTÁL, Zdeněk. *Lineární algebra* [online]. 2000 [cit. 2007-12-14]. Dostupný z WWW: <[http://www.am.vsb.cz/vondrak/education/LA\\_IT/Books/LINALG.pdf](http://www.am.vsb.cz/vondrak/education/LA_IT/Books/LINALG.pdf)>.
- [5] ŠOCHMAN, Jan, PAJDLA, Tomáš. *Matematický model kamery* [online]. 06. 03. 2003 [cit. 2007-12-14]. Dostupný z WWW: <<ftp://cmp.felk.cvut.cz/pub/cmp/articles/sochman/Sochman-TR-2002-11.pdf>>.
- [6] HARRIS, Chris, STEPHENS, Mike. *A combined corner and edge detector* [online]. 1988 [cit. 2007-12-18]. Dostupný z WWW: <[http://www.csse.uwa.edu.au/~pk/research/matlabfns/Spatial/Docs/Harris/A\\_Combined\\_Corner\\_and\\_Edge\\_Detector.pdf](http://www.csse.uwa.edu.au/~pk/research/matlabfns/Spatial/Docs/Harris/A_Combined_Corner_and_Edge_Detector.pdf)>.
- [7] WERNER, Tomáš. *Harris corner detector* [online]. 2007 [cit. 2007-12-18]. Dostupný z WWW: <<http://cmp.felk.cvut.cz/cmp/courses/PVR/2007/LabsZ/docs/harris.pdf>>.
- [8] LOWE, David. *Distinctive Image Features from Scale-Invariant Keypoints* [online]. 2004 [cit. 2007-12-18]. Dostupný z WWW: <<http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>>.
- [9] BROWN, Matthew, LOWE, David. *Automatic Panoramic Image Stitching using Invariant Features* [online]. 2006 [cit. 2007-12-18]. Dostupný z WWW: <<http://www.cs.ubc.ca/~mbrown/papers/ijcv2006.pdf>>.
- [10] HESS, Rob. *Rob Hess School of EECS @ Oregon State University* [online]. 2007 [cit. 2008-05-14]. Dostupný z WWW: <<http://web.engr.oregonstate.edu/~hess/index.html>>.
- [11] *The GNU General Public License* [online]. 2008 [cit. 2008-05-14]. Dostupný z WWW: <<http://www.gnu.org/licenses/gpl.html>>.
- [12] ŘÍHA, Kamil. *Pokročilé techniky zpracování obrazu*. [s.l.] : [s.n.], 2007. 105 s.

- [13] HORČÍK, Rostislav. *Taylorovy řady* [online]. 2004 , 24.5.2004 [cit. 2008-05-26].  
Dostupný z WWW: <<http://www2.cs.cas.cz/~horcik/Teaching/complex6.pdf>>.
- [14] *Pořizování panoramatických fotografií - Nodal point* [online]. [2005] , 26.2.2006  
[cit. 2008-05-26]. Dostupný z WWW:  
<[http://www.makofoto.cz/nikon/tech\\_podpora/nodal\\_point.htm](http://www.makofoto.cz/nikon/tech_podpora/nodal_point.htm)>.

## A. Obsah přiloženého CD

### [CD-ROM]

- └ [Diplomová práce] - text diplomové práce a metadata práce ve formátu PDF
- └ [Testovací snímky] - zdrojové snímky použité v kapitole 5.3
- └ [Zdrojové kódy] - zdrojové kódy aplikace
- └ [Zkompilovaný program] - přeložená spustitelná verze aplikace